Using dimension reduction to improve outbreak predictability of multistrain diseases

Leah B. Shaw[1], Lora Billings[2], and Ira B. Schwartz[1]

[1]*Naval Research Laboratory, Plasma Physics Division, Nonlinear Systems Dynamics Section, Code 6792, Washington, DC 20375*

[2]*Department of Mathematical Sciences, Montclair State University, Upper Montclair, NJ 07043*

Abstract

Multistrain diseases have multiple distinct coexisting serotypes (strains). For some diseases, such as dengue fever, the serotypes interact by antibody-dependent enhancement (ADE), in which infection with a single serotype is asymptomatic, but contact with a second serotype leads to higher viral load and greater infectivity. We present and analyze a dynamic compartmental model for multiple serotypes exhibiting ADE. Using center manifold techniques, we show how the dynamics rapidly collapses to a lower dimensional system. Using the constructed reduced model, we can explain previously observed synchrony between certain classes of primary and secondary infectives [14]. Additionally, we show numerically that the center manifold equations apply even to noisy systems. Both deterministic and stochastic versions of the model enable prediction of asymptomatic individuals that are difficult to track during an epidemic. We also show how this technique may be applicable to other multistrain disease models, such as those with cross-immunity.

Key words: center manifold analysis, epidemic models, multistrain disease, dengue

1 Introduction

In the study of infectious disease, a problem of interest is the coexistence of multiple strains. Examples of persistent co-circulating multistrain diseases include influenza [1], malaria [12], and dengue fever [11]. More recently, the avian flu viruses, including H5N1, were reported to coexist in several genotypes until 2004 [17]. Other examples of viruses possessing multiple strains may be seen in corona viruses, such as severe acute respiratory syndrome (SARS) [18]. The presence of multiple strains adds greater complication to models of disease

dynamics due to an increasing number of stages through different infection-recovery combinatorics. Reducing the dimensionality of the model is desirable, both to obtain a simpler model and to understand how the strains interact.

In this paper, we focus on dengue fever, which has four co-circulating serotypes. It is believed that following infection with and recovery from one serotype, cross-reactive antibodies act to enhance the infectiousness of a subsequent infection by another serotype [16]. This phenomenon is termed antibody-dependent enhancement (ADE). Primary infections are less severe, often asymptomatic, but secondary infections are associated with more serious illness and greater risk for dengue hemorrhagic fever (DHF).

Multistrain diseases with ADE-induced dynamics have been modeled previously [10,14,5]. Models of multistrain steady state endemic behavior and its stability were considered for two strains in [9], where both mosquito population and humans were included. Although ADE was not explicitly modeled, the authors did consider different parameter values of contact susceptibility. Values of contact susceptibility less than unity model cross-immunity, whereas values greater than unity model increased susceptibility due to secondary infection. Ferguson *et al.* modeled two serotypes of dengue and showed that ADE can lead to oscillations and chaotic behavior [10]. Schwartz *et al.* developed a similar model with non-overlapping compartments for greater ease of interpretation [14]. The bifurcation structure was obtained for both the autonomous model and a model with seasonal forcing. Chaotic solutions were found for a wide range of parameter values. It was observed from numerical simulations that outbreaks of the four serotypes could occur asynchronously but that certain primary and secondary infective compartments remained synchronized. In particular, all compartments currently infected with a particular strain exhibited in-phase outbreaks. Cummings *et al.*, using the same model, explored the competitive advantage that ADE confers to viruses. Other models for dengue fever (e.g., [8]) have included the mosquito vector as in [9] but have not included the ADE effect.

Several models of multistrain diseases with cross-immunity rather than ADE have appeared in the literature [1,6,8]. In these models, individuals who have recovered from infection with one serotype have reduced susceptibility to infection with other serotypes. Oscillatory dynamics have been observed.

Summarizing the results to date, in models of secondary infections where either ADE or cross-immunity is considered, the steady state equilibrium is not always stable, and persistent oscillations may occur. For models with ADE, although ADE is the cause of oscillation onset, more complicated chaotic or stochastic behavior is observed over a much wider range of parameter values than those of periodic oscillations [14]. Currently, there is no theory for predicting the interrelationship between primary and secondary infective com-

2

partments.

In this paper, we consider the problem of revealing the dynamical structure between the secondary and primary infectives. Specifically, we analyze the model presented [14] and demonstrate that the dynamics can be reduced to a lower dimensional model. Our analysis motivates the in-phase behavior previously reported for primary and secondary infectives. The lower dimensional system allows the prediction of primary infective levels from knowledge of other compartment values. This result may be useful in disease monitoring because little epidemiological data exists for the asymptomatic primary infectives. We show numerically that our predictions hold even in noisy systems. We also discuss the relevance of this technique to other multistrain diseases that do not display ADE, such as influenza, in which infection with one strain yields partial immunity to other strains.

2   Description of general n-serotype model

We study the model of [14] for n co-circulating serotypes. An individual can contract each serotype only once and is assumed to gain immunity to all serotypes after infection from two distinct serotypes. This compartmental model divides the population into disjoint sets and follows the size of these compartments as a percentage of the whole population over time. The variable definitions are as follows:

> s Susceptible to all serotypes;
>
> x_i Primary infectious with serotype i
>
> r_i Primary recovered from serotype i
>
> x_{ij} Secondary infectious, currently infected with serotype j,
>
> but previously had i $(i \neq j)$.

The model is a system of $n^2 + n + 1$ ordinary differential equations describing the rates of change of the population within each compartment:

$$\frac{ds}{dt} = \mu - \beta s \sum_{i=1}^{n} \left(x_i + \phi_i \sum_{j \neq i} x_{ji} \right) - \mu_d s \tag{1}$$

$$\frac{dx_i}{dt} = \beta s \left(x_i + \phi_i \sum_{j \neq i} x_{ji} \right) - \sigma x_i - \mu_d x_i \tag{2}$$

3

$$\frac{dr_i}{dt} = \sigma x_i - \beta r_i \sum_{j \neq i} \left(x_j + \phi_j \sum_{k \neq j} x_{kj} \right) - \mu_d r_i \qquad (3)$$

$$\frac{dx_{ij}}{dt} = \beta r_i \left(x_j + \phi_j \sum_{k \neq j} x_{kj} \right) - \sigma x_{ij} - \mu_d x_{ij}. \qquad (4)$$

The parameters μ, μ_d, β, and σ denote birth, death, contact, and recovery rates, respectively. Antibody-dependent enhancement is governed by the parameters ϕ_i. Rates of infection due to primary infectives are of the form $\beta s x_i$, as in a standard SIR model, while infection rates due to secondary infectives are weighted by the appropriate ϕ, taking the form $\beta \phi_j s x_{ij}$. When $\phi = 1$, there is no ADE, and both primary and secondary infectives are equally infectious. When $\phi > 1$, the ADE appears as an enhancement factor in the nonlinear terms involving secondary infectives.

Although the contact rate β could be time dependent (e.g., due to seasonal fluctuations in the mosquito vector population), we assume it to be constant here for simplicity. The fixed parameters throughout the paper are given by: $\mu = 0.02$, $\mu_d = 0$ or $\mu_d = 0.02$, $\beta = 400$, and $\sigma = 100$, all with units of $years^{-1}$. These values are consistent with estimates used previously in modeling dengue [5,2]. We assume all serotypes have equal ADE factors: $\phi_i = \phi$ for all i. The ADE factor has not been measured from epidemiological data, so we test our results for various ADE values.

The dynamics of Eq. 1-4 have been studied previously [14,2]. A bifurcation diagram in the ADE factor ϕ is given in Fig. 1 for the four serotype model. The dynamics are qualitatively similar for other values of n, the number of serotypes. For $\phi \in [1, \phi_c)$, the system has a stable endemic steady state. At a critical ϕ value, ϕ_c, the system undergoes a Hopf bifurcation and begins to oscillate periodically. For slightly larger ϕ values, the periodic solutions become unstable and the system oscillates chaotically. Chaotic oscillations persist for most larger values for ϕ, with the exception of narrow windows of stable periodic solutions.

It should be noted that previous studies of this model [14,2] used a death rate equal to the birth rate ($\mu_d = \mu = 0.02$) so that the population remained constant in time. However, the endemic steady state cannot easily be obtained analytically when the mortality terms of Eqs. 1-4 are included. In our analysis, we omit the small mortality terms by setting $\mu_d = 0$. The dynamics with $\mu_d = 0$ are qualitatively similar to that with $\mu_d = 0.02$ and have the same bifurcation structure. We demonstrate numerically that our results hold well even for the system with mortality.

The larger parameters β, σ may be scaled by the small parameter μ, defining
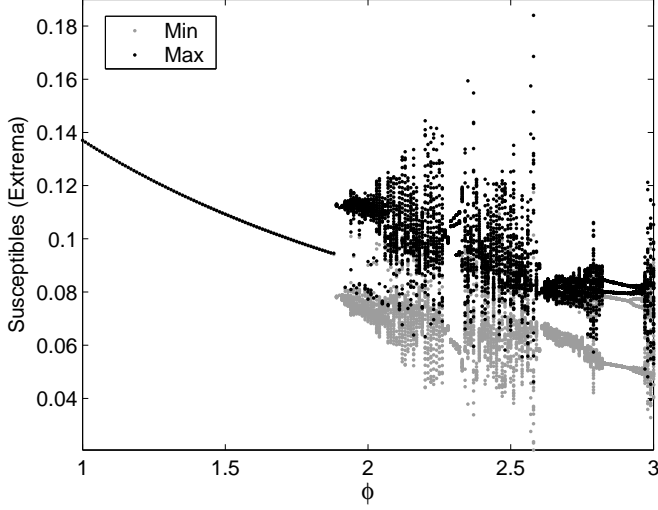
$$\beta = \beta_0/\mu, \sigma = \sigma_0/\mu, \qquad (5)$$

Fig. 1. A bifurcation diagram for Eqs. 1-4 with $n = 4$, $\beta = 400$, $\sigma = 100$, $\mu = 0.02$, and $\mu_d = 0.02$. Shown is a Poincaré section constructed from the extrema of a time series for susceptibles of length $t = 100$ years after a transient of $t = 5000$ years. The minima are denoted by grey points and the maxima are overlaid in black points. The Hopf bifurcation occurs at approximately $\phi = 1.88333$.

where β_0, σ_0 are $\mathcal{O}(1)$. Eqs. 1-4 with $\mu_d = 0$ have an endemic steady state of

$$\left(s_0 = \frac{\sigma_0}{\beta_0(\phi + 1)}, x_{i,0} = \frac{\mu^2}{n\sigma_0}, r_{i,0} = \frac{\sigma_0}{\beta_0(n-1)(\phi+1)}, x_{ij,0} = \frac{\mu^2}{n(n-1)\sigma_0} \right) \tag{6}$$

for all i, j.

3 Center manifold analysis

Although the full n-serotype model has $n^2 + n + 1$ dimensions, the attractor lies in a much lower dimensional space. We claim that the attractor actually lies in $2n + 1$ dimensions to a good approximation, due to a relationship between primary and secondary infectives. This result can be shown using center manifold analysis. We now outline a derivation for the case of two serotypes and state the center manifold equations for several other general cases of interest.

3.1 $n = 2$ serotypes

We begin with Eqs. 1-4 for two serotypes with no mortality ($\mu_d = 0$). We shift the variables so that the endemic fixed point (Eq. 6) is at the origin:

5

$\bar{s} = s - s_0, \bar{x}_i = x_i - x_{i,0}, \bar{r}_i = r_i - r_{i,0}, \bar{x}_{ij} = x_{ij} - x_{ij,0}$. The Jacobian in the shifted variables, evaluated at the origin, is given by

$$\frac{1}{1+\phi} \begin{bmatrix} 0 & -\sigma_0 & -\sigma_0 & 0 & 0 & -\phi\sigma_0 & -\phi\sigma_0 \\ 0 & -\phi\sigma_0 & 0 & 0 & 0 & \phi\sigma_0 & 0 \\ 0 & 0 & -\phi\sigma_0 & 0 & 0 & 0 & \phi\sigma_0 \\ 0 & \sigma_0(1+\phi) & -\sigma_0 & 0 & 0 & 0 & -\phi\sigma_0 \\ 0 & -\sigma_0 & \sigma_0(1+\phi) & 0 & 0 & -\phi\sigma_0 & 0 \\ 0 & \sigma_0 & 0 & 0 & 0 & -\sigma_0 & 0 \\ 0 & 0 & \sigma_0 & 0 & 0 & 0 & -\sigma_0 \end{bmatrix} \tag{7}$$

to zeroth order in μ.

The Jacobian is not diagonalizable, as it has only 5 linearly independent eigenvectors. However, a new set of variables w is defined as follows:

$$w = \frac{1}{1+\phi}\left[\bar{x}_{21} - \bar{x}_1, \bar{x}_{12} - \bar{x}_2, (1+\phi)\bar{s}, \bar{x}_1 + \phi\bar{x}_{21}, \bar{x}_2 + \phi\bar{x}_{12},\right.$$
$$\left.\phi(\bar{x}_{21} - \bar{x}_1) - (1+\phi)\bar{r}_1, \phi(\bar{x}_{12} - \bar{x}_2) - (1+\phi)\bar{r}_2\right] \tag{8}$$

The transformation matrix for this change of variables contains the 5 linearly independent eigenvectors of the Jacobian and 2 additional vectors selected to be linearly independent and put the system in the desired form.

In the new variables, rescaling time to $\tau = t/\mu$, the dynamics are described by

$$\frac{dw_1}{d\tau} = -\sigma_0 w_1 + \beta_0 w_4(w_7 + \phi w_2 - w_3) + \frac{\mu^2\beta_0}{2\sigma_0}(w_7 + \phi w_2 - w_3) \tag{9}$$

$$\frac{dw_2}{d\tau} = -\sigma_0 w_2 + \beta_0 w_5(w_6 + \phi w_1 - w_3) + \frac{\mu^2\beta_0}{2\sigma_0}(w_6 + \phi w_1 - w_3) \tag{10}$$

$$\frac{dw_3}{d\tau} = -\sigma_0(w_4 + w_5) - \beta_0(1+\phi)w_3(w_4 + w_5) - \frac{\mu^2\beta_0}{\sigma_0}(1+\phi)w_3 \tag{11}$$

$$\frac{dw_4}{d\tau} = \beta_0 w_4(\phi^2 w_2 + \phi w_7 + w_3) + \frac{\mu^2\beta_0}{2\sigma_0}(\phi^2 w_2 + \phi w_7 + w_3) \tag{12}$$

$$\frac{dw_5}{d\tau} = \beta_0 w_5(\phi^2 w_1 + \phi w_6 + w_3) + \frac{\mu^2\beta_0}{2\sigma_0}(\phi^2 w_1 + \phi w_6 + w_3) \tag{13}$$

$$\frac{dw_6}{d\tau} = \sigma_0(w_4 - w_5) + \beta_0 w_4(-\phi^2 w_2 + \phi w_3 - \phi w_7)$$

$$+ \beta_0 w_5 \left[-\phi(1+\phi)w_1 - (1+\phi)w_6 \right]$$

$$+ \frac{\mu^2 \beta_0}{2\sigma_0} \left[-\phi^2 w_2 + \phi w_3 - \phi w_7 - \phi(1+\phi)w_1 - (1+\phi)w_6 \right] \tag{14}$$

$$\frac{dw_7}{d\tau} = \sigma_0(w_5 - w_4) + \beta_0 w_4 \left[-\phi(1+\phi)w_2 - (1+\phi)w_7 \right]$$

$$+ \beta_0 w_5(-\phi^2 w_1 + \phi w_3 - \phi w_6)$$

$$+ \frac{\mu^2 \beta_0}{2\sigma_0} \left[-\phi(1+\phi)w_2 - (1+\phi)w_7 - \phi^2 w_1 + \phi w_3 - \phi w_6 \right] \tag{15}$$

$$\frac{d\mu}{d\tau} = 0 \tag{16}$$

The Jacobian of Eqs. 9-15, evaluated at the origin and to zeroth order in μ, is

$$\begin{bmatrix}
-\sigma_0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\sigma_0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\sigma_0 & -\sigma_0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \sigma_0 & -\sigma_0 & 0 & 0 \\
0 & 0 & 0 & -\sigma_0 & \sigma_0 & 0 & 0
\end{bmatrix} \tag{17}$$

Eqs. 9-16 may be written in the compact form

$$\dot{\mathbf{u}} = \mathbf{A}\mathbf{u} + \mathbf{f}(\mathbf{u}, \mathbf{v}, \mu)$$
$$\dot{\mathbf{v}} = \mathbf{B}\mathbf{v} + \mathbf{g}(\mathbf{u}, \mathbf{v}, \mu)$$
$$\dot{\mu} = 0$$

where $\mathbf{u} = (w_1, w_2)$, $\mathbf{v} = (w_3, \ldots, w_7)$, where $\mathbf{A}$ and $\mathbf{B}$ are constant matrices such that the eigenvalues of $\mathbf{A}$ have negative real parts and the eigenvalues of $\mathbf{B}$ have zero real parts, and $\mathbf{f}$ and $\mathbf{g}$ are second order in $\mathbf{u}, \mathbf{v}, \mu$. Therefore, center manifold theory [3] can be applied. The system will rapidly collapse onto a lower dimensional manifold defined by

$$w_1 = h(\mu, w_3, w_4, \ldots, w_7) \tag{18}$$
$$w_2 = l(\mu, w_3, w_4, \ldots, w_7) \tag{19}$$

We expand h and l to second order

$$w_1 \approx h_0 + \sum_{i=3}^{7} h_i w_i + h_\mu \mu + \sum_{i=3}^{7} \sum_{j=i}^{7} h_{i,j} w_i w_j + \sum_{i=3}^{7} h_{\mu,i} \mu w_i \qquad (20)$$

$$w_2 \approx l_0 + \sum_{i=3}^{7} l_i w_i + l_\mu \mu + \sum_{i=3}^{7} \sum_{j=i}^{7} l_{i,j} w_i w_j + \sum_{i=3}^{7} l_{\mu,i} \mu w_i. \qquad (21)$$

To solve for the coefficients of the expansion, we write equalities for $\frac{dw_1}{d\tau}$ and $\frac{dw_2}{d\tau}$. This is done by equating the right hand sides of Eqs. 9-10 with the time derivatives of Eqs. 20-21, using Eqs. 11-15 to substitute for $\frac{dw_3}{d\tau}, \ldots, \frac{dw_7}{d\tau}$ and Eqs. 20-21 to substitute for w_1, w_2 as needed. The coefficients of the expansion can be obtained by equating coefficients of like powers.

The coefficients $h_0, h_i, h_\mu, h_{\mu,i}$ and $l_0, l_i, l_\mu, l_{\mu,i}$ are all 0, while some of the $h_{i,j}$ and $l_{i,j}$ are nonzero. After carrying out the above program, Eqs. 20-21 simplify to

$$w_1 = -\frac{\beta_0}{\sigma_0} w_3 w_4 - 2\frac{\beta_0}{\sigma_0} w_4 w_5 + \frac{\beta_0}{\sigma_0} w_4 w_7 \qquad (22)$$

$$w_2 = -\frac{\beta_0}{\sigma_0} w_3 w_5 - 2\frac{\beta_0}{\sigma_0} w_4 w_5 + \frac{\beta_0}{\sigma_0} w_5 w_6 \qquad (23)$$

when the coefficients are substituted.

Finally, we rewrite Eqs. 22-23 in the original variables. Thus we obtain the following approximation for the invariant manifold onto which the system collapses:

$$\sigma_0(\bar{x}_1 - \bar{x}_{21}) = \beta_0 \left[\bar{s} - \bar{r}_2 + \frac{2-\phi}{1+\phi}\bar{x}_2 + \frac{3\phi}{1+\phi}\bar{x}_{12} \right] (\bar{x}_1 + \phi\bar{x}_{21})$$

$$\sigma_0(\bar{x}_2 - \bar{x}_{12}) = \beta_0 \left[\bar{s} - \bar{r}_1 + \frac{2-\phi}{1+\phi}\bar{x}_1 + \frac{3\phi}{1+\phi}\bar{x}_{21} \right] (\bar{x}_2 + \phi\bar{x}_{12}) \qquad (24)$$

The above equations for the center manifold are our main result of the paper. To bring out more of the structure, we make the following observations. We generally observed in numerical simulations that the infective compartments (and their deviations from the fixed point) are small compared to the susceptibles and recovereds. Thus the infective correction terms to $\bar{s} - \bar{r}_i$ may be dropped to obtain a simpler expression for the center manifold:

$$\sigma_0(\bar{x}_1 - \bar{x}_{21}) = \beta_0 (\bar{s} - \bar{r}_2) (\bar{x}_1 + \phi\bar{x}_{21})$$
$$\sigma_0 (\bar{x}_2 - \bar{x}_{12}) = \beta_0(\bar{s} - \bar{r}_1)(\bar{x}_2 + \phi\bar{x}_{12}) \qquad (25)$$

Given values for the susceptibles, primary recovereds, and secondary infectives, the primary infective compartments may be computed from the approx-
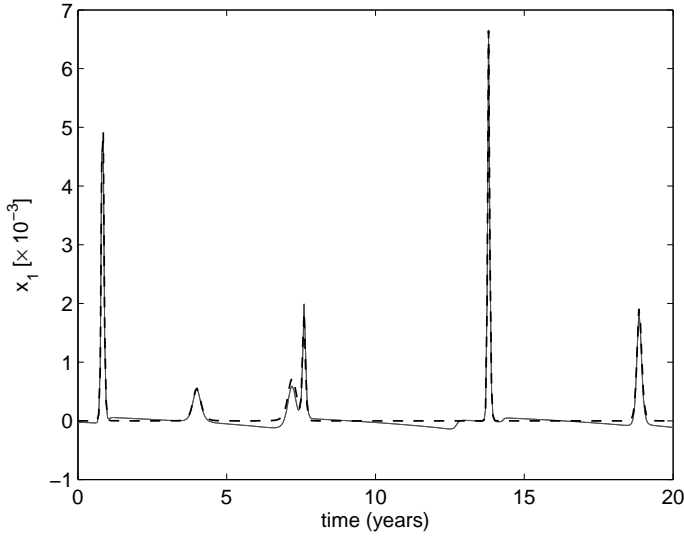
Fig. 2. Primary infectives with strain 1 vs. time for two-strain model with no mortality. Solid gray line: prediction from approximate center manifold equations (predictions from Eqs. 24 and 25 lie on same curve); dashed black line: time trace from direct integration. The curves from Eqs. 24 and 25 coincide almost exactly. The dynamics are chaotic with the parameter values used here: $\beta = 400$, $\sigma = 100$, $\mu = 0.02$, $\mu_d = 0$, $\phi = 2.8$.

imate center manifold equations (either Eqs. 24 or 25). Figure 2 compares the center manifold prediction for primary infectives with that from simulations. (Results from Eqs. 24 and 25 are nearly indistinguishable.) The second order approximation to the center manifold leads to reasonable predictions for the primary infectives. In particular, we accurately predict the time and approximate magnitude of the bursts, which correspond to epidemic outbreaks. The prediction for the infectives is occasionally negative because the center manifold equations are for the $\bar{x}_i$, the deviations of infectives from their fixed point, and sometimes the predicted deviations are large enough that adding the fixed point $x_{i,0}$ yields negative values for the infective variables x_i. However, these errors are not physically important because they occur in quiescent regions where epidemic outbreaks are not occurring. Moreover, the timing of the predicted outbreaks agrees well with actual outbreak time of the full system.

3.1.1 Asymmetric ADE factors

Antibody enhancement factors have not been measured experimentally, and it is possible that different serotypes could have different levels of enhancement. We address the possibility of unequal ADE factors for the two serotype case.

A procedure similar to that used with symmetric ADE factors may be followed. In this case, an expansion for the location of the fixed point must also be used.
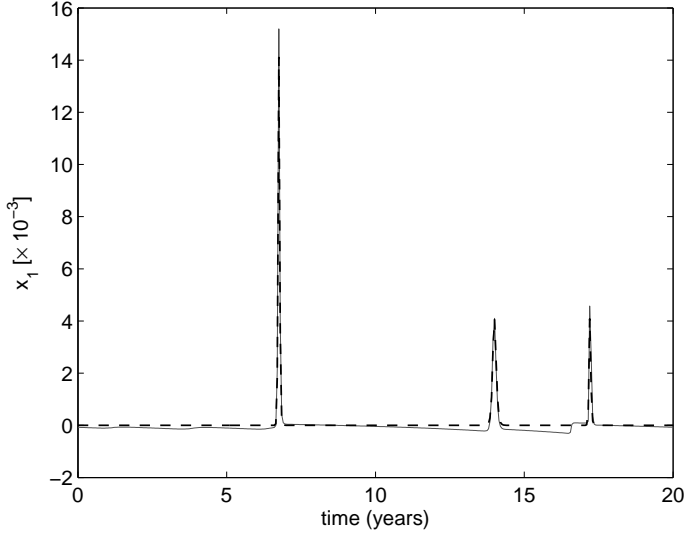
Fig. 3. Primary infectives with strain 1 vs. time for two-strain model with no mortality and asymmetric ADE factors. Solid gray line: prediction from approximate center manifold equations (Eq. 26 and 27); dashed black line: time trace from direct integration. Parameters: $\beta = 400$, $\sigma = 100$, $\mu = 0.02$, $\mu_d = 0$, $\phi_1 = 2.5$, $\phi_2 = 3.75$. The ADE asymmetry measure is $\epsilon = 0.5$.

For two strains with $\phi_1 = \phi$, $\phi_2 = \phi(1 + \epsilon)$ and $|\epsilon| \ll 1$, the center manifold is approximated (to second order in the variables and μ, ϵ) by

$$\sigma_0 \left[\bar{x}_1 - \bar{x}_{21} + \epsilon \left(\bar{x}_1 + \phi \bar{x}_{21} \right) \frac{\sigma_0}{2 \left(\sigma_0 + 1 \right) \left(1 + \phi \right)} \right]$$
$$= \beta_0 \left[\bar{s} - \bar{r}_2 + \frac{2 - \phi}{1 + \phi} \bar{x}_2 + \frac{3\phi}{1 + \phi} \bar{x}_{12} \right] \left(\bar{x}_1 + \phi \bar{x}_{21} \right) \tag{26}$$

$$\sigma_0 \left[\bar{x}_2 - \bar{x}_{12} - \epsilon \left(\bar{x}_2 + \phi \bar{x}_{12} \right) \frac{\sigma_0}{2 \left(\sigma_0 + 1 \right) \left(1 + \phi \right)} \right]$$
$$= \beta_0 \left[\bar{s} - \bar{r}_1 + \frac{2 - \phi}{1 + \phi} \bar{x}_1 + \frac{3\phi}{1 + \phi} \bar{x}_{21} \right] \left(\bar{x}_2 + \phi \bar{x}_{12} \right) \tag{27}$$

The center manifold equations again allow fairly accurate predictions of the primary infective compartments for small ϵ (data not shown). Even when the $|\epsilon| \ll 1$ assumption does not hold, the time and amplitude of bursts in the primary infectives are generally predicted accurately (Fig. 3).

3.2 $n > 2$ strains

It is important to extend the analysis to a larger number of serotypes. In particular, there are four observed serotypes for dengue fever. The center manifold

10

technique may again be applied for $n > 2$ strains. In the expressions for the center manifold, it is convenient to define the sum of secondary infectives currently infected with strain k:

$$\bar{z}_k = \sum_{i=1, i \neq k}^{n} \bar{x}_{ik} \tag{28}$$

For dengue fever, primary infectives are generally asymptomatic, and most hospital cases are secondary infectives. The current infecting serotype can be determined from serology measurements [13]. $\bar{z}_k$, the sum of secondary infectives with serotype k, may be proportional to the number of hospital cases with serotype k. This quantity may be relevant to disease monitoring in addition to providing a shorthand notation for expressing the center manifold equations.

We have completed center manifold analysis for $n = 2, 3, 4$ serotypes. The following equations for the center manifold summarize our results for $n = 2, 3, 4$ and extrapolate them for general n:

$$\sigma_0 \left[\bar{x}_k - \bar{z}_k \right] = \beta_0 \left[\bar{s} - \sum_{i \neq k} \bar{r}_i + f_k(\bar{x}_p, \bar{x}_{pq}, n, \phi) \right] (\bar{x}_k + \phi \bar{z}_k) \tag{29}$$

$$\sigma_0 \left[(n-1)\bar{x}_{jk} - \bar{z}_k \right] = \beta_0 \left[(n-1)\bar{r}_j - \sum_{i \neq k} \bar{r}_i + g_{jk}(\bar{x}_p, \bar{x}_{pq}, n, \phi) \right] (\bar{x}_k + \phi \bar{z}_k) \tag{30}$$

where

$$f_k = \frac{n - (n-1)\phi}{(n-1)(1+\phi)} \left(\sum_{i \neq k} \bar{x}_i \right) + \frac{(2n-1)\phi}{(n-1)(1+\phi)} \left(\sum_{i,l,i \neq l} \bar{x}_{il} - \sum_{i \neq k} \bar{x}_{ik} \right) \tag{31}$$

$$g_{jk} = \frac{n - (n-1)\phi}{(n-1)(1+\phi)} \left(\sum_{i \neq k} \bar{x}_i - (n-1)\bar{x}_j \right)$$

$$+ \frac{(2n-1)\phi}{(n-1)(1+\phi)} \left(\sum_{i,l,i \neq l} \bar{x}_{il} - \sum_{i \neq k} \bar{x}_{ik} - (n-1) \sum_{i \neq j} \bar{x}_{ij} \right) \tag{32}$$

We observe that the terms f_k and g_{jk} may be neglected. They are linear functions of primary and secondary infective compartments. At the fixed point, the infective compartments are order μ, while the susceptibles and recovereds are order 1 (Eq. 6). It therefore is reasonable that deviations of the infectives from the fixed point, represented by the variables $\bar{x}_p, \bar{x}_{pq}$, would generally be small compared to $\bar{s}, \bar{r}_i$. This assumption generally holds true in our simulations.

For each of the n strains, Eqs. 29 and 30 provide $n - 1$ linearly independent equations, since some of the Eqs. 30 are linearly dependent, allowing $n(n-1)$ dimensions to be eliminated. Thus the dynamics of the system have dimension $2n + 1$. If the susceptible and primary recovered compartments are known, a single quantity $\bar{z}_k$ (Eq. 28) for each strain is sufficient to describe the infective dynamics for that strain. Primary and secondary infectives can then easily be computed using Eqs. 29 and 30. We find that similar accuracy is attained when the f_k, g_{jk} are set to 0 for convenience as with the f_k, g_{jk} terms retained.

Figure 4 compares the predictions from the center manifold equations (Eqs. 29 and 30) with actual values for sample primary and secondary infectives. The full four-strain model with mortality ($\mu_d = 0.02$) was used in the simulations. Predictions were made by assuming that susceptibles, primary recovereds, and sums of secondary infectives ($\bar{z}_k$) were known. Although the derivation of the center manifold equations omitted mortality terms, they generate accurate predictions for the full model.

3.2.1 Phase synchrony between compartments

It has been observed in numerical simulations that primary and secondary infectives currently infected with the same serotype oscillate in phase synchrony [14]. Based on the center manifold reduction in dimension, we provide some analysis to motivate this effect.

We convert the shifted variables $(\bar{s}, \bar{x}_i, \bar{r}_i, \bar{x}_{ij})$ to the original variables (s, x_i, r_i, x_{ij}) and use the center manifold equations 29-30 (with f, g terms omitted as explained above) for substitution in the model ordinary differential equations, Eqs. 2 and 4. We obtain

$$\frac{dx_k}{dt} = -\sigma z_k + \beta \sum_{i \neq k} r_i \left(x_k + \phi z_k \right) + \beta \left(s - \sum_{i \neq k} r_i \right) \left(x_{k,0} + \phi z_{k,0} \right) \quad (33)$$

$$(n-1)\frac{dx_{jk}}{dt} = -\sigma z_k + \beta \sum_{i \neq k} r_i \left(x_k + \phi z_k \right)$$

$$+ \beta \left((n-1)r_j - \sum_{i \neq k} r_i \right) \left(x_{k,0} + \phi z_{k,0} \right) \quad (34)$$

where we have defined

$$z_k = \sum_{i=1, i \neq k}^{n} x_{ik} \quad (35)$$
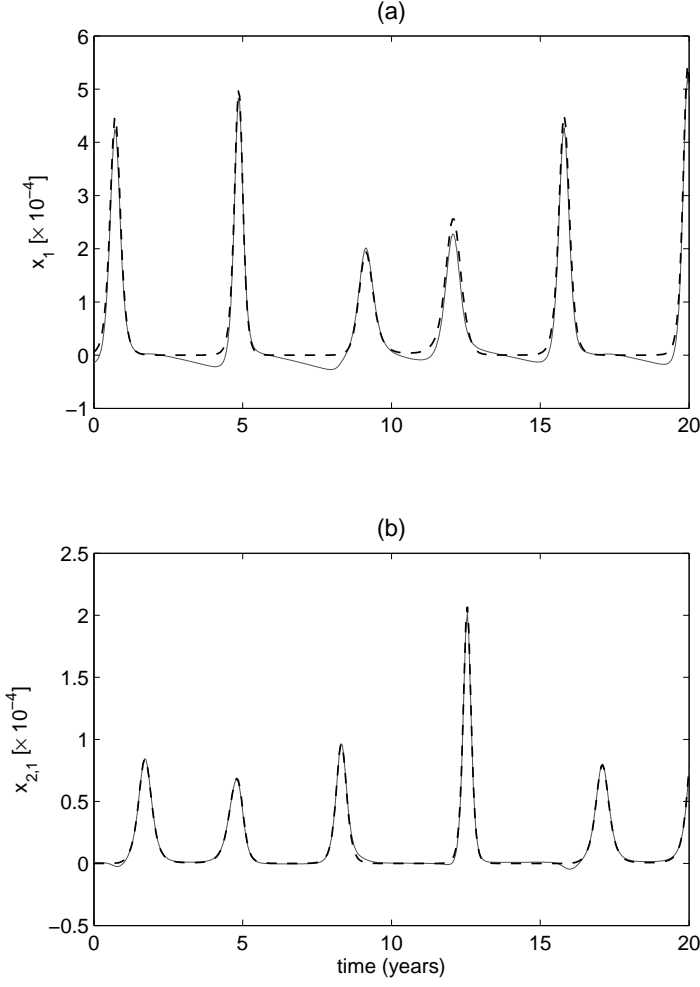
for convenience.

Fig. 4. (a) Primary infectives with strain 1 vs. time for four-strain model with mortality; (b) secondary infectives contracting strain 1 then 2 vs. time for four-strain model with mortality. Solid gray lines: prediction from approximate center manifold equations (Eqs. 29 and 30 with f_k, g_{jk} terms neglected); dashed black lines: time traces from direct integration. The dynamics are chaotic with the parameter values used here: $\beta = 400$, $\sigma = 100$, $\mu = 0.02$, $\mu_d = 0.02$, $\phi = 2.0$.

Taking the difference between Eqs. 33 and 34 and converting back to rescaled variables yields

$$\frac{d\bar{x}_k}{dt} - (n-1)\frac{d\bar{x}_{jk}}{dt} = \beta \left[\bar{s} - (n-1)\bar{r}_j\right](x_{k,0} + \phi z_{k,0}) \tag{36}$$

The difference between Eqs. 29 and 30 provides a useful substitution relation:

$$\sigma \left[\bar{x}_k - (n-1)\bar{x}_{jk}\right] = \beta \left[\bar{s} - (n-1)\bar{r}_j\right](\bar{x}_k + \phi \bar{z}_k) \tag{37}$$

13

Substitution of Eq. 37 into 36 gives an equation for how the difference between primary and secondary infectives evolves in time:

$$\frac{d}{dt}\left[\bar{x}_k(t) - (n-1)\bar{x}_{jk}(t)\right] = K\frac{\bar{x}_k(t) - (n-1)\bar{x}_{jk}(t)}{\bar{x}_k(t) + \phi\bar{z}_k(t)} \tag{38}$$

where $K = \sigma(x_{k,0} + \phi z_{k,0})$ is a positive constant of order μ. We restrict our consideration of Eq. 38 to those intervals on which $\bar{x}_k(t) + \phi\bar{z}_k(t)$ is negative, thus avoiding singularities in $[\bar{x}_k(t) + \phi\bar{z}_k(t)]^{-1}$. That is, we consider intervals in which the sum of primary and secondary infectives with serotype k (weighted by the ADE factor) is below the steady state value. In the chaotic regime, these intervals occur during dropouts when serotype k is present in low levels. Such intervals cover the majority of the time domain and are interrupted by outbreaks of serotype k. (Cf. Figure 4.)

Integrating Eq. 38, we find that

$$\bar{x}_k(t) - (n-1)\bar{x}_{jk}(t)$$
$$= [\bar{x}_k(t_0) - (n-1)\bar{x}_{jk}(t_0)]\exp\left\{K\int_{t_0}^{t}[\bar{x}_k(s) + \phi\bar{z}_k(s)]^{-1}\,ds\right\} \tag{39}$$

From our previous assumption, the argument of the integral is negative and is of order $1/\mu^2$, since the infectives range between 0 and the negative of the fixed point. Therefore, $|\bar{x}_k(t) - (n-1)\bar{x}_{jk}(t)| \ll 1$, and $\bar{x}_k(t)$ and $(n-1)\bar{x}_{jk}(t)$ approach each other rapidly, with their difference decreasing with $\exp[-\mathcal{O}(1/\mu)]$, where $|\mu| \ll 1$.

We have demonstrated that primary and secondary infectives with serotype k are phase synchronized, differing only by an exponentially small term, during the intervals in which levels of serotype k are low. Outbreaks in the infectives occur in bursts, so although the relation $|\bar{x}_k(t) - (n-1)\bar{x}_{jk}(t)| \ll 1$ is not expected to hold during outbreaks, the outbreaks have fast time scale dynamics. Therefore the duration of the outbreak is short enough that phase synchrony is not lost. It may be possible to examine the phase locking during the outbreaks, but that would require a detailed multiscale analysis, which is beyond the scope of the present paper. Even so, our results help to explain the phase synchrony previously observed between primary and secondary infectives currently infected with the same strain [14].

14

4   Stochastic perturbations

We next consider Eqs. 1-4 with multiplicative noise and discuss whether Eqs. 29 and 30 can be applied to a noisy system. We include the noise in our numerical integration as follows.

Let $\mathbf{Y} = \{s, x_i, r_i, x_{ij}\}_{i,j=1,...,n}$ and $\dot{\mathbf{Y}} = \mathbf{F}(\mathbf{Y})$. Define natural log coordinates $\mathbf{y} = \{y_i\}$, where $\ln(Y_i) = y_i$ and $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$. Standard change of coordinates relates the two systems by:

$$\dot{Y}_i = Y_i f_i(\ln \mathbf{Y}), \quad \text{since} \quad \frac{\dot{Y}_i}{Y_i} = \dot{y}_i \tag{40}$$

We add noise to the natural log coordinates:

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}) + \eta. \tag{41}$$

Here, η is a vector of random variables with mean zero and standard deviation σ_n. Transforming back to the original phase space, we obtain

$$\dot{Y}_i = Y_i f_i(\ln \mathbf{Y}) + \eta_i Y_i \tag{42}$$

$$\dot{Y}_i = F_i(\mathbf{Y}) + \eta_i Y_i \tag{43}$$

Therefore, the noise is multiplicative and scaled by each component. Eq. 41 was integrated numerically using a stochastic Euler method.

The predictions from Eqs. 29 and 30 hold approximately true for a wide range of noise standard deviations σ_n. Sample results are shown in Figure 5. For the parameter values in Figure 5, the endemic steady state is stable for $\sigma_n = 0$. Adding noise perturbs the system away from the steady state, and as the noise standard deviation σ_n increases, the amplitude of the disease outbreaks increases.

To quantify the accuracy of the center manifold prediction for various σ_n, we define a metric for the "distance" of a trajectory from the center manifold:

$$d = \left\langle \sum_{i=1}^{n} |x_i(t) - x_{i,pred}(t)| \right\rangle_t, \tag{44}$$

where $x_{i,pred}$ is the predicted primary infective value using Eqs. 29 and 30 (neglecting f_k, g_{jk} terms) and $\langle \cdot \rangle_t$ denotes a time average. We compute d by sampling every 0.01 year for 10^4 years (after first running for 10^4 years to
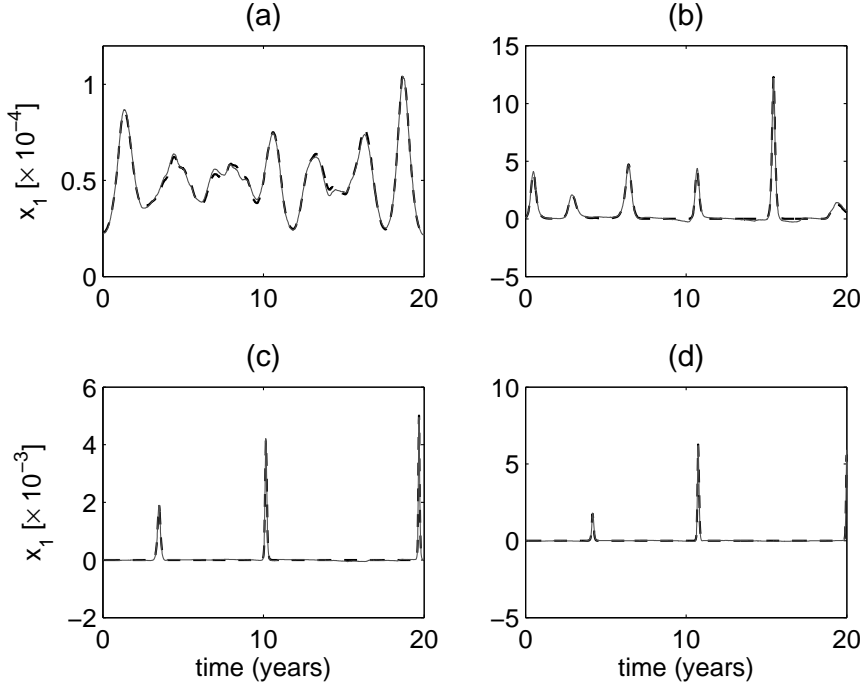
15

Fig. 5. Primary infectives with strain 1 vs. time for noisy four-strain model with no mortality, for various noise standard deviations. Solid gray lines: prediction from approximate center manifold equations (Eqs. 29 and 30 with f_k, g_{jk} terms neglected); dashed black lines: time traces from direct integration. (a) $\sigma_n = 0.01$, (b) $\sigma_n = 0.05$, (c) $\sigma_n = 0.1$, (d) $\sigma_n = 0.15$. Other parameter values: $\beta = 400$, $\sigma = 100$, $\mu = 0.02$, $\mu_d = 0$, $\phi = 1.9$.

remove transients) in order to obtain good statistics. Sample results are given in Figure 6(a). The "distance" d from the center manifold, i.e., the error in the center manifold prediction, increases as the noise increases. However, the amplitude of the outbreaks also increases with larger noise, so it is not unexpected that d would increase.

We additionally compute a normalized distance in lieu of a percent error so that the effect of noise may be more carefully assessed. (The actual percent error is not used because it would be greatly elevated by large percent errors in the dropout periods when the infectives are low.) The 10^4 year sampling period is divided into windows of 100 years, and the maximum value of the primary infective x_1 (arbitrarily chosen) is computed for each window. These maxima give a scale for the outbreak amplitudes for each σ_n. The distance d is then divided by the average maximum for each σ_n to compute a normalized distance from the center manifold. Results are given in Figure 6(b). Using this metric, we find that the normalized error in the center manifold prediction actually decreases in noisy systems.
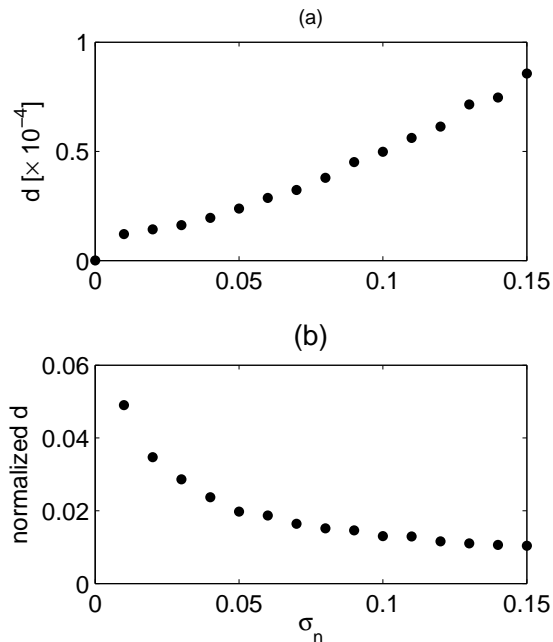
Fig. 6. (a) Average "distance" d from center manifold vs. noise standard deviation; (b) normalized distance vs. noise standard deviation, where distance and normalized d are as defined in the text. Parameters: $\beta = 400$, $\sigma = 100$, $\mu = 0.02$, $\mu_d = 0$, $\phi = 1.9$.

5   Conclusions and discussion

We have analyzed a model for multistrain diseases with antibody-dependent enhancement. For a disease with n strains and two sequential infections, the model has n^2+n+1 equations. Using center manifold analysis, we have reduced the model to $2n + 1$ dimensions. Although we have derived an approximation to the manifold on which the solutions for both deterministic and stochastic systems lie, if we use this approximation and evolve the dynamics of a model constrained to the manifold, we do not necessarily see the same bifurcation structure. This is due to the fact that the analysis itself is local and done near the steady state endemic point. Despite that limitation, the lower dimensional system has other uses which we have described here.

We have partially explained the synchrony between primary and secondary infective compartments currently infected with the same serotype. During the intervals when a serotype is present in low levels, primary and secondary infectives of that serotype are approximately related by a constant of proportionality. Although our derivation does not hold during outbreaks of that serotype, the outbreaks occur in short bursts during which phase synchrony is not lost. As mentioned above, a full multiscale analytic treatment may lend more insight into the phase relationship between primary and secondary infections of a given serotype during the outbreak periods.

17

Our center manifold approximation enables prediction of each primary and secondary infective value if the disease parameters, susceptible and recovered compartments, and sums of secondary infectives (Eq. 28) are known. The prediction matches well with numerical simulations even when significant amounts of noise are added to the system. It has been observed [13] that most hospital cases of dengue fever are secondary infections. From serotype measurements of these hospital cases, estimation of the number of secondary infectives that have a particular serotype i should be possible. On the other hand, there is little epidemiological data on the frequency of primary infections because patients with primary infections tend not to be as seriously ill and are not admitted to hospitals. They are typically considered asymptomatic. Obtaining primary infective data for dengue monitoring purposes might require random sampling of large numbers of apparently healthy individuals, which would not be feasible economically nor practically. Our center manifold equations might be used to improve monitoring of dengue outbreaks by predicting compartments that cannot readily be measured, such as the primary infectives.

The fact that susceptible and recovered compartment data is necessary to make predictions for primary infectives is a potential drawback. However, we observe in numerical simulations that the susceptibles and recovereds vary more slowly than the infective compartments, which quickly spike during an outbreak. Therefore, it is possible that susceptible and recovered data could be obtained by sampling a population less frequently than would be required for infective data. If such sampling is possible, the center manifold equations will yield information about primary infectives from indirect measurements from population data.

The applicability of our method to other models is also of interest. Multistrain models for diseases with cross-immunity, in which recovery from one strain leads to reduced susceptibility to further infection with other strains, have been developed [4,1,6,8]. We have analyzed a two-strain model with cross-immunity [15]:

$$\frac{ds}{dt} = \mu - \beta s \sum_{i=1}^{n} \left(x_i + \sum_{j \neq i} x_{ji} \right) - \mu_d s \tag{45}$$

$$\frac{dx_i}{dt} = \beta s \left(x_i + \sum_{j \neq i} x_{ji} \right) - \sigma x_i - \mu_d x_i \tag{46}$$

$$\frac{dr_i}{dt} = \sigma x_i - \psi \beta r_i \sum_{j \neq i} \left(x_j + \sum_{k \neq j} x_{kj} \right) - \mu_d r_i \tag{47}$$

$$\frac{dx_{ij}}{dt} = \psi \beta r_i \left(x_j + \sum_{k \neq j} x_{kj} \right) - \sigma x_{ij} - \mu_d x_{ij}, \tag{48}$$

where $\psi < 1$ expresses the degree of cross-immunity and $n = 2$. This two-strain cross-immunity model is introduced in [4]. This model differs from the ADE model in that rather than giving the secondary infectives a different infectivity (weighted by ϕ), we give the primary recovereds a different susceptibility to further infection (weighted by ψ) than the susceptible compartment. Using the coordinate transformation given in Eq. 8 (setting $\phi = 1$), we can reduce Eqs. 45-48 from 7 to 5 equations and show a relationship between primary and secondary infectives. The eigenvalues of the Jacobian corresponding to rapidly collapsing directions are $n(n-1)$ eigenvalues equal to $-\sigma$, the recovery rate. The center manifold analysis holds because the recovery rate is rapid compared to the birth rate. It seems reasonable to expect that other models with cross-immunity might also be reduced and a relationship between primary, secondary, and perhaps higher order infectives might be derived. Because cross-immunity models are often written in a different form with overlapping compartments (e.g.,[1,6,8]), some reformulation of the equations may be necessary before a center manifold approach can be used.

The results presented here were primarily for a symmetric multistrain model with equal disease parameters for each strain. However, the case of unequal ADE factors was briefly considered, and a center manifold was found when the difference between ADE factors was a small parameter (Eqs. 26-27). To model a multistrain disease with realistic parameters, it may be necessary to relax the symmetry constraint and derive results for an asymmetric system. Such a derivation is unwieldy because the endemic steady state can no longer be determined analytically (except as an approximate expansion). New techniques may be necessary to extend our current approach to more asymmetric models.

Another potentially interesting extension is to include spatial spread of the disease. It has been observed from epidemiological data in Thailand that while dengue outbreaks within a single town may be of primarily one serotype, concurrent outbreaks in another town may have a different serotype predominant [7]. A similar center manifold analysis might be attempted on a patch model, in which the population is divided into several spatially distinct patches that are coupled.

References

[1] V. Andreasen, J. Lin, and S. A. Levin. The dynamics of cocirculating influenza strains conferring partial cross-immunity. J. Math. Bio., 35:825–842, 1997.

[2] L. Billings, I. B. Schwartz, L. B. Shaw, and M. McCrary. In preparation.

[3] J. Carr. Applications of centre manifold theory. Springer-Verlag, New York, 1981.

[4] C. Castillo-Chavez, H. W. Hethcote, Andreasen V., Levin S. A., and W. M. Liu. J. Math. Biol., 27:233–258, 1989.

[5] D. A. T. Cummings, I. B. Schwartz, L. Billings, L. B. Shaw, and D. S. Burke. Proc. Natl. Acad. Sci. USA, 102:15259–15264, 2005.

[6] J.H.P. Dawes and J.R. Gog. The onset of oscillatory dynamics in models of multistrain diseases. J. Math. Bio., 45:471–510, 2002.

[7] T. P. Endy, A. Nisalak, S. Chunsuttiwat, D. H. Libraty, S. Green, A. L. Rothman, D. W. Vaughn, and F. E. Ennis. Am. J. Epidemiology, 156:52–59, 2002.

[8] L. Esteva and C. Vargas. Coexistence of different serotypes of dengue virus. J. Math. Bio., 46:31–47, 2003.

[9] Z. Feng and J. X. Velasco-Hernandez. Competitive exclusion in a vector host model for the denuge fever. J. Math. Bio., 35:523–544, 1997.

[10] N. Ferguson, R. Anderson, and S. Gupta. The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. Proc. Natl. Acad. Sci. USA, 96:790–794, 1999.

[11] N. M. Ferguson, C. A. Donnelly, and R. M. Anderson. Transmission dynamics and epidemiology of dengue: insights from age-stratified sero-prevalence surveys. Proc. R. Soc. London, Ser. B, 354:757–768, 1999.

[12] S. Gupta, K. Trenholme, R. M. Anderson, and K. P. Day. Antigenic diversity and the transmission dynamics of plasmodium-falciparum. Science, 263:961–963, 1994.

[13] A. Nisalak, T. P. Endy, S. Nimmannitya, S. Kalayanarooj, U. Thisayakorn, R. M. Scott, D. S. Burke, C. H. Hoke, B. L. Innis, and D. W. Vaughn. Serotype-specific dengue virus circulation and dengue disease in bangkok, thailand from 1973 to 1999. Am. J. Trop. Med. Hyg., 68:191–202, 2003.

[14] I. B. Schwartz, L. B. Shaw, D. A. T. Cummings, L. Billings, M. McCrary, and D. S. Burke. Phys. Rev. E, 72:066201, 2005.

[15] L. B. Shaw. Unpublished.

[16] D. W. Vaughn, S. Green, S. Kalayanarooj, B. L. Innis, and et al. Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J. Infect. Diseases, 181:2, 2000.

[17] R. G. Webster and D. J. Hulse. Microbial adaptation and change:avian influenza. Rev. Sci. Tech. Off. Int. Epiz., 23:453–465, 2004.

[18] Z.-Y. Yang, H. C. Werner, W. P. Kong, K. Leung, and et al. PNAS, 102:797–801, 2005.